HAL
open science

Apport des fluences catégorielles dans le diagnostic du mild cognitive impairment : revue systématique de littérature

Jeanne Arminjon-Lupatelli, Élena Vaccaro

To cite this version:

Jeanne Arminjon-Lupatelli, Élena Vaccaro. Apport des fluences catégorielles dans le diagnostic du mild cognitive impairment : revue systématique de littérature. Sciences cognitives. 2020. dumas-02948206

HAL Id: dumas-02948206

https://dumas.ccsd.cnrs.fr/dumas-02948206v1

Submitted on 24 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ACADÉMIE DE PARIS

SORBONNE UNIVERSITÉ

MÉMOIRE POUR LE CERTIFICAT DE CAPACITÉ D'ORTHOPHONISTE

APPORT DES FLUENCES CATÉGORIELLES DANS LE DIAGNOSTIC DU *MILD COGNITIVE IMPAIRMENT :* REVUE SYSTÉMATIQUE DE LITTÉRATURE

NOM DU DIRECTEUR DE MÉMOIRE :

Agnès WEILL-CHOUNLAMOUNTRY, orthophoniste

ANNÉE UNIVERSITAIRE 2019-2020

ARMINJON-LUPATELLI Jeanne

jeanne.lupatelli@gmail.com

VACCARO Elena

vaccaro.elena@gmail.com

Affichage des mails autorisé par les auteurs

Remerciements communs

Merci à Agnès Weill-Chounlamountry d'avoir accepté de conduire ce travail et de l'avoir soutenu dans ses multiples rebondissements. Merci à Anne-Laure Cortadellas-Doucet, notre rapporteur, pour son soutien bienveillant. Merci à Marie Villain pour sa relecture attentive. Merci à Anaïs Migeot et Sophie Janicaud pour leurs conseils.

Merci à Eric Kendjo et Thierry Ancelle, biostatisticiens, dont l'aide a été précieuse pour valider notre démarche statistique.

Remerciements (Jeanne Arminjon-Lupatelli)

Cinq années pour franchir le Rubicon… par le pont d'Austerlitz ! Entre la gare de Lyon et la Pitié-Salpêtrière, Dijon et Paris, journalisme et orthophonie.

Un immense merci à Jean, mon mari, de m'avoir encouragée, épaulée, supportée… alors que je m'investissais parfois sans mesure dans ce nouveau métier.

Merci à nos enfants Marie, Louis et Rita d'avoir gardé le sourire en voyant leur maman partir dès l'aurore vers Paris et en lui demandant le soir si elle avait « gagné » !

Merci à mes parents et beaux-parents pour leur indéfectible soutien moral et logistique.

Merci à mes maîtres de stage pour leurs généreux passages de flambeaux.

Merci à Elena, notre duo m'a été précieux pour aller au bout de cette reconversion.

Remerciements (Elena Vaccaro)

Alors que la ligne d'arrivée - qui est aussi celle d'un nouveau départ - de ce marathon de cinq ans est proche, je me retourne un instant pour apprécier le chemin accompli : indéniablement sinueux, mais surtout le plus riche que j'ai eu à parcourir. Je regarde en avant et il me tarde de poursuivre, notamment auprès des patients que j'aime tant rencontrer.

Je voudrais remercier mes proches, famille et amis, pour leur soutien sans faille pendant ces cinq années. Merci à eux d'avoir toujours cru en moi et de m'avoir encouragée si affectueusement.

Merci à mes maîtres de stage pour leur patience et tout ce qu'elles m'ont appris.

Un merci très spécial pour Jeanne, bien plus que ma « binôme de mémoire », une amie, avec qui j'ai tant aimé discuter et travailler tout au long de cette formation.

Merci aussi à mes autres amies de promotion, qui ont apporté de la légèreté à ces années studieuses.

Résumé

Le concept de *mild cognitive impairment* (MCI) ou trouble neurocognitif léger se définit comme une phase de transition entre le vieillissement normal et l'entrée dans une pathologie neurodégénérative. Le diagnostic précoce des patients MCI est un enjeu majeur de santé publique afin d'anticiper une prise en charge adaptée.

Cette revue systématique de littérature vise à déterminer si l'épreuve de fluence catégorielle (FC) est un outil pertinent pour le diagnostic précoce du MCI en comparant les résultats obtenus à cette épreuve par les sujets MCI et contrôles.

À l'issue d'un processus de sélection rigoureux, 14 articles publiés ces dix dernières années et issus des banques de données PubMed, PsycNET et Cochrane Library, ont été inclus dans cette étude. L'analyse statistique réalisée sur RevMan5 a permis de calculer une grandeur d'effet à partir des données extraites des études incluses (4675 participants, dont 2446 MCI et 2229 contrôles) et de conclure à une différence significative entre les moyennes aux FC « animaux » des groupes MCI et contrôles.

Cette revue systématique a confirmé la pertinence de l'épreuve de fluence catégorielle dans le cadre du diagnostic précoce du MCI et a étudié son intérêt par rapport au MMSE (Mini Mental State Examination), test presque systématiquement proposé au cours du diagnostic des maladies neurodégénératives. La passation de cette tâche rapide et facile à administrer pourrait être systématisée en cas de suspicion d'un trouble neurocognitif léger.

Mots clés : trouble neurocognitif léger, fluences verbales, fluences catégorielles, évaluation, diagnostic

Abstract

Mild cognitive impairment (MCI) is defined as a transitional phase between normal aging and the onset of a neurodegenerative pathology. Early diagnosis of MCI patients is a critical public health issue in order to anticipate an appropriate treatment.

The aim of this systematic review was to determine whether the semantic fluency task is a relevant tool for early diagnosis of MCI, by comparing the scores obtained by MCI subjects to those of controls.

Through a rigorous selection process, fourteen articles published in the last ten years and identified on PubMed, PsycNET and Cochrane Library databases, were included in this study. The statistical analysis performed on RevMan5 has highlighted an effect size from the data collected in the fourteen studies (4675 participants, including 2446 MCI and 2229 controls). A significant difference in mean scores was found between MCI and control groups for the "animal" semantic fluency task.

This review confirmed the relevance of such a task in early diagnosis of MCI and investigated its value in relation to MMSE (Mini Mental State Examination), an almost systematically proposed test for the diagnosis of neurodegenerative diseases. The semantic fluency task, quick and easy to administer, could be generalized whenever a mild neurocognitive disorder is suspected.

Keywords: mild cognitive impairment, verbal fluency, semantic fluency, assessment, diagnosis

Attestations de non-plagiat

Je soussignée JEANNE ARMINJON-LUPATELLI déclare être pleinement consciente que le plagiat de documents ou d'une partie d'un document publiés sur toutes formes de support, y compris l'Internet, constitue une violation des droits d'auteur ainsi qu'une fraude caractérisée. En conséquence, je m'engage à citer toutes les sources que j'ai utilisées pour écrire ce mémoire.

Signature :

Je soussignée ELENA VACCARO déclare être pleinement consciente que le plagiat de documents ou d'une partie d'un document publiés sur toutes formes de support, y compris l'Internet, constitue une violation des droits d'auteur ainsi qu'une fraude caractérisée. En conséquence, je m'engage à citer toutes les sources que j'ai utilisées pour écrire ce mémoire.

Signature :

Table des illustrations

Figures

Figure A : Diagramme PRISMA (diagramme de flux ou *flow chart*)..... 9
Figure B : *Forest plot* des scores à l'épreuve des FC pour les 14 études incluses..... 12
Figure C : *Forest plot* des scores à l'épreuve du MMSE pour 12 études..... 13
Figure D : Biais de publication représenté sous forme de *funnel plot* 16

Tableaux

Tableau 1 : Étiologies probables des sous-types de MCI..... 3
Tableau 2 : Impact des variables démographiques sur les performances en fluences verbales chez le sujet sain..... 4
Tableau 3 : Indicateurs de qualité des études incluses 10
Tableau 4 : Synthèse des données démographiques des études incluses 11
Tableau 5 : Comparaison des FC au MMSE par le calcul du « d de Cohen » 14
Tableau 6 : Synthèse par étude des types de MCI et de démences 14

Liste des abréviations

FC : fluence catégorielle

FP : fluence phonémique

FV : fluence verbale

MA : maladie d'Alzheimer

MCI : *mild cognitive impairment*

NC : sujets contrôles

Introduction (Elena Vaccaro)

Le concept de *mild cognitive impairment* (MCI) ou trouble neurocognitif léger se définit comme une phase de transition entre le vieillissement normal et l'entrée dans une pathologie neurodégénérative. Il est caractérisé par quatre critères : modification cognitive subjective par rapport à une situation antérieure ; performances altérées dans un ou plusieurs domaines cognitifs ; autonomie préservée en vie quotidienne ; pas de signe de démence (Tangalos et Petersen, 2018).

Le diagnostic précoce des patients MCI est un enjeu majeur de santé publique. Il permet d'anticiper une prise en charge adaptée et donc de maintenir les capacités fonctionnelles, ralentir le déclin cognitif et l'entrée en institution (Hwang, Boes, Nyffeler et Schuepfer, 2019). Il a été montré qu'une évaluation centrée uniquement sur la mémoire épisodique était peu spécifique, dans la mesure où son altération atteint également le vieillissement normal (Kim *et al.*, 2019). Les connaissances sémantiques restent quant à elles préservées dans le vieillissement sain alors qu'elles se dégradent de façon précoce en cas de pathologie neurodégénérative (Kim *et al.*, 2019). L'exploration de la mémoire sémantique pourrait donc contribuer à poser le diagnostic de MCI (Lehrner *et al.,* 2017 ; Mortamais *et al.*, 2017), notamment grâce aux épreuves de fluences verbales (FV) qui seraient un marqueur efficace pour discriminer le vieillissement normal et l'entrée dans un processus neurodégénératif. Elles présentent l'avantage de s'adapter à toutes les langues et cultures et de ne nécessiter que quelques minutes de passation (Venneri, Jahn-Carta, De Marco, Quaranta et Marra, 2018). Si les fluences phonémiques (FP) consistent à générer le plus de mots possible commençant par une lettre donnée en temps limité, les fluences catégorielles (FC) consistent, quant à elles, à produire le plus de mots possible appartenant à une catégorie sémantique donnée en temps limité (60 sec. ou 2 min. en général pour les deux types de fluences). Chez les patients présentant la maladie d'Alzheimer (MA), il a été décrit une plus grande vulnérabilité des FC par rapport aux FP en lien avec l'évolution neurophysiopathologique de la maladie (Mortamais *et al.,* 2017 ; Venneri *et al.,* 2018).

L'objectif de cette revue systématique de la littérature est de déterminer si les FC sont discriminantes dans le diagnostic du MCI. Nous supposons que les FC des sujets MCI sont significativement déficitaires par rapport à celles des sujets contrôles et discutons du bénéfice de l'ajout d'une épreuve de FC dans l'évaluation initiale afin de dépister plus finement et/ou plus précocement les troubles neurocognitifs légers que le test Mini Mental State Examination (MMSE), *gold standard* pratiqué presque systématiquement.

Matériel et méthode (Elena Vaccaro)

Afin de répondre à l'objectif principal de notre étude qui était de déterminer l'apport des fluences catégorielles dans le diagnostic du MCI, nous avons mené une revue systématique de littérature, de type méta-analyse. Ce type de revue de littérature permet de faire la synthèse de la recherche sur une question précise par une méthodologie d'analyse critique systématique, appuyée par des données statistiques. Elle consiste à identifier, sélectionner, évaluer et synthétiser les études en lien avec cette question qui ont un niveau de preuve élevé et une méthode rigoureuse explicitée, puis à en réaliser l'analyse critique (Bettany-Saltikov, 2010a).

Banques de données et équations de recherche

Pour atteindre la plus grande exhaustivité possible, l'identification des articles pertinents s'est faite à travers trois bases de données : MEDLINE/PubMed, Cochrane Library et APA PsycINFO/PsycNET.

PubMed est l'interface de la base de données MEDLINE donnant accès à plus de 30 millions de références bibliographiques dans les domaines de la biologie et de la médecine, essentiellement en langue anglaise. Pour cette recherche, l'équation suivante a été utilisée : « verbal fluency » (Title/Abstract) AND « mild cognitive impairment » (Title/Abstract) NOT « intervention » (Title/Abstract) NOT « training » (Title/Abstract), pour la période allant de 2010 à 2020. Nous avons en effet choisi de limiter notre requête aux travaux les plus récents, parus depuis 2010. Cette équation de recherche a donné un résultat de 196 articles (29/02/2020).

La Cochrane Library regroupe trois bases de données spécialisées en médecine et en santé : Cochrane Database of Systematic Reviews (CDSR), Cochrane central Register of Controlled Trials-Clinical Trials (CENTRAL) et Cochrane Clinical Answers (CCAs). Sa méthodologie rigoureuse de sélection des revues et sa mise à jour fréquente en font une des sources de données probantes les plus fiables. L'équation de recherche utilisée sur Cochrane Library était la suivante : « verbal fluency » (Title/Abstract/Keyword) AND « mild cognitive impairment » (Title/Abstract/Keyword) NOT « intervention » (Title/Abstract/Keyword), pour la période allant de 2010 à 2020. Cette recherche a donné un résultat de 22 articles dans CENTRAL (29/02/2020).

PsycINFO est la base de données de l'American Psychological Association (APA) référençant des articles dans le domaine de la psychologie et des disciplines connexes (médecine, psychiatrie, sciences de l'éducation, sociologie, etc.). Nous avons utilisé l'équation de recherche suivante sur l'interface PsycNET : « verbal fluency » (Abstract)

AND « mild cognitive impairment » (Abstract), pour la période de 2010 à 2020. Avec cette recherche, 6 résultats ont été obtenus (29/02/2020).

Protocole et critères d'inclusion/exclusion

Avant de commencer le travail de recension et d'analyse, un protocole a été établi pour définir la stratégie de lecture des articles et ainsi limiter les biais. Des critères d'inclusion et d'exclusion des études ont été définis en fonction de l'objectif principal de notre recherche : déterminer s'il existe une différence significative entre les scores à la tâche de FC « animaux » (donner le plus de noms d'animaux possible en 1 minute) des sujets contrôles et des patients MCI. Si une telle différence pouvait être montrée et renforcée par l'obtention d'un effet global sur une population importante (cumul de la population des études incluses), nous pourrions alors conclure que ce type de fluence présente un intérêt avéré dans le diagnostic du MCI.

Le concept de MCI

Rappelons la variété de la présentation clinique du MCI, qui n'est plus aujourd'hui associé seulement à la MA au stade préclinique. Pour tenir compte de cette diversité, des critères plus larges, connus sous le nom de « consensus de Stockholm », ont été admis et prévalent actuellement dans la définition du MCI (Winblad *et al.*, 2004). Ainsi, le concept de MCI s'est enrichi de quatre principaux sous-types, en lien avec des étiologies probables (Tableau 1), le rendant plus précis mais aussi potentiellement plus hétérogène selon le ou les sous-types étudiés au sein de chaque étude :

- aMCI-sd (*amnestic MCI single domain*) : trouble mnésique seulement
- aMCI-md (*amnestic MCI multiple domains*) : trouble mnésique et un ou plusieurs autres troubles cognitifs
- naMCI-sd (*non amnestic MCI single domain*) : un seul trouble cognitif non mnésique
- naMCI-md (*non amnestic MCI multiple domains*) : plusieurs troubles cognitifs non mnésiques

Tableau 1. Étiologies probables des sous-types de MCI.

	Dégénérative	**Vasculaire**	**Psychiatrique**
aMCI-sd	MA	-	DP
aMCI-md	MA	DV	DP
naMCI-sd	DFT ou DCL	DV	-
naMCI-md	DCL	-	-

Tableau réalisé d'après Petersen, 2004 ; Winblad *et al.*, 2004.
MA : maladie d'Alzheimer. DFT : démence fronto-temporale. DCL : démence à corps de Lewy. DV : démence vasculaire. DP : dépression.

Petersen lui-même, qui est à l'origine du concept de MCI, met en garde les cliniciens et chercheurs sur la variabilité des résultats trouvés dans la littérature dans le champ d'étude du MCI (Petersen, 2004). Il détaille trois principales sources de confusion :

- La multiplicité des échelles d'évaluation et des tests neuropsychologiques employés et les différents scores-seuils choisis.
- La définition du vieillissement « normal » qui exclut en général les comorbidités (pas de prise de psychotropes, pas de maladie neurologique ou psychiatrique, etc.) s'avère très restrictive et peu en phase avec la réalité du vieillissement.
- Le recrutement des patients dans des cliniques ou centres mémoire accroît la probabilité d'être en présence de troubles dégénératifs.

Les variables à contrôler pour les épreuves de FC

Une étude de normalisation des épreuves de fluences verbales impliquant 1133 sujets a conclu que les variables qui ont un impact sur les scores obtenus aux fluences catégorielles (dont animaux) sont l'âge, l'éducation, le métier exercé ou exercé par le passé et, dans une moindre mesure, le sexe. « Les variables associées avec une meilleure performance en fluence catégorielle sont un âge plus jeune, un niveau d'éducation plus élevé, davantage d'occupations intellectuelles par le passé, pas de dépression et le fait d'être une femme. » (Auriacombe, Fabrigoule, Lafont, Jacqmin-Gadda et Dartigues, 2001 – Tableau 2).

L'impact du sexe sur les scores aux fluences catégorielles est modéré avec des performances légèrement supérieures pour les femmes, alors qu'il est négligeable concernant les fluences phonémiques.

L'âge influence très fortement les performances aux FC. Chaque année supplémentaire infléchit le score global aux FC de 0,4 point.

L'impact du niveau d'éducation et du métier sur ces scores est fort, mais pas autant que sur les fluences phonémiques encore plus dépendantes du niveau académique, de la richesse du vocabulaire et de la profession exercée (Auriacombe *et al.*, 2001).

Tableau 2. Impact des variables démographiques sur les performances en fluences verbales chez le sujet sain (d'après Auriacombe *et al.*, 2001).
Pas d'impact : - ; léger impact **:** +/- impact notable : + ; impact fort : ++ ; impact très fort : +++

Variables démographiques	Fluences catégorielles (FC)	Fluences phonémiques (FP)
sexe	+	-
âge	+++	+++
éducation (années d'études)	++	+++
métier	++	+++

Critères d'inclusion et d'exclusion

Les critères d'inclusion ont été les suivants :

- articles publiés entre janvier 2010 et février 2020
- en anglais ou en français
- publication dans une revue avec comité de lecture
- un groupe MCI et un groupe contrôle
- une épreuve de FC « animaux » en 60 secondes (la plus répandue dans la littérature)
- l'âge et le niveau d'éducation moyens des participants

Les critères d'exclusion ont été les suivants :

- études centrées sur la maladie de Parkinson ou pathologie autre que la MA
- études centrées sur la remédiation plutôt que sur le diagnostic
- études centrées sur des données biologiques et/ou anatomiques
- études pour lesquelles les données chiffrées n'étaient pas exploitables

Sélection des articles

Une fois ces critères définis, il nous a été possible, à la lecture des titres et des résumés de tous les articles (n=208), de déterminer lesquels seraient retenus. Ce travail a été réalisé par une double lecture, démarche préconisée dans le cadre des revues systématiques de littérature pour valider la sélection des articles et contrôler les erreurs potentielles d'inattention ou d'interprétation (Maison, 2010). Un protocole de sélection rigoureux et structuré est aussi un gage de reproductibilité (Bettany-Saltikov, 2010a).

Lors de la mise en commun, les études faisant l'objet d'un accord initial (n=202) ont été sélectionnées pour une lecture complète ou exclues. Les cas de discordance (n=6), ont été discutés pour parvenir à un accord.

Collecte des données

L'extraction des données s'est faite à travers trois grilles principales. Un premier tableau de synthèse incluant les 208 études de départ (doublons exclus) indiquait pour chacune l'inclusion ou l'exclusion ainsi que le motif et l'étape d'exclusion.

Un second tableau (voir synthèse en annexe), concernant les études incluses, comprenait :

- Des données descriptives :
    - population et recrutement
    - critères diagnostiques

    - démarche expérimentale
    - résultats et conclusions
    - points forts et limites de l'étude
- Des données quantitatives pour chaque groupe :
    - nombre de sujets
    - âge moyen
    - niveau d'éducation moyen
    - score moyen et écart-type aux FC « animaux »
    - score moyen et écart-type au MMSE

Enfin, le travail de revue de littérature implique de vérifier la qualité méthodologique des études incluses sans quoi les interprétations des résultats peuvent être faussées (Zaugg, Savoldelli, Sabatier et Durieux, 2014). En ce sens, nous avons complété un troisième tableau afin d'évaluer la qualité des études incluses (Tableau 3).

Qualité des études

Robustesse des essais cliniques

Apprécier la fiabilité des études revient à croiser deux principaux critères : le nombre de participants et le degré de contrôle expérimental (Ebbels, 2017). Ainsi, une étude avec beaucoup de participants et un design expérimental rigoureux mérite la confiance du clinicien et peut être considérée comme robuste dans le cadre d'une revue systématique de littérature. Inversement, les conclusions d'une étude peu contrôlée et avec peu de participants doivent être considérées avec prudence. Ebbels (2017) relève quatre combinaisons possibles de ces deux critères afin d'évaluer la robustesse des études, allant du plus au moins robuste :

- B : contrôle expérimental élevé, nombre de participants élevé
- A : contrôle expérimental élevé, nombre de participants faible
- D : contrôle expérimental faible, nombre de participants élevé
- C : contrôle expérimental faible, nombre de participants faible

Dix modèles expérimentaux sont définis : du moins contrôlé (1 : expérience clinique et anecdotes) au plus contrôlé (10 : essai randomisé contrôlé). Cette classification, élaborée à l'origine pour apprécier la robustesse des études interventionnelles, peut s'adapter aux études diagnostiques. Concernant celles de notre revue de littérature, le degré de contrôle expérimental est assez élevé et très homogène : pour les 14 études incluses, il se situe au niveau 8 : étude de groupes avec comparaison de participants non randomisée. En

revanche, nous constatons une variabilité conséquente par rapport au nombre de participants : de 52 à 1820, ce qui situe les études incluses au niveau A ou B.

Indicateurs de qualité

Au-delà de la robustesse telle que la définit Ebbels (2017), un ensemble d'indicateurs nous ont semblé pertinents pour apprécier la qualité globale des études retenues. À notre connaissance, aucun standard de qualité n'a été publié pour évaluer les études à visée diagnostique dans le champ des maladies neurodégénératives. C'est pourquoi, nous avons dû définir notre propre liste d'indicateurs (Tableau 3), adaptée de démarches existantes, souvent conçues pour valider des études interventionnelles et/ou issues d'autres champs de connaissance. Nos principales sources ont été les suivantes :

- Gersten *et al.* (2005), Quality Indicators for Group Experimental and Quasi-Experimental Research in Special Education.
- The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analysis (cité par Bettany-Saltikov, 2010b).
- Critical Appraisal Skills Programme (2006), Oxford Centre for Triple Value Healthcare (cité par Bettany-Saltikov, 2010b).

Dans un premier temps, ces indicateurs ont été appréciés par deux examinateurs. Un accord inter-évaluateur a été trouvé pour les rares éléments de discordance.

Analyse statistique

Une analyse a été menée à l'aide du logiciel RevMan5, outil proposé par The Cochrane Collaboration pour le traitement des revues de littérature. L'objectif de ce traitement statistique était de savoir d'une part si la différence entre les scores moyens à la tâche FC « animaux » obtenus par les sujets contrôles et les sujets MCI était significative, et d'autre part si la différence entre les scores moyens au MMSE obtenus par les deux groupes était significative. Ces différences de moyennes entre les deux groupes (z-score) pour chacun des deux tests étudiés (FC et MMSE) représentent les grandeurs d'effet.

Pour comparer ces deux tests entre eux, un « d de Cohen » a été calculé. Il s'agit d'une mesure descriptive utilisée pour estimer une taille d'effet entre deux moyennes (effet faible, moyen ou fort).

Le logiciel RevMan5 nous a également permis d'obtenir un score d'hétérogénéité des études incluses et d'évaluer le risque de biais de publication, deux éléments incontournables de toute revue de littérature. L'indice I^2 fournit ainsi une estimation du pourcentage de variabilité due à l'hétérogénéité (Zaugg *et al.*, 2014). Le risque de biais de

publication a été appréhendé sous la forme d'une représentation graphique (*funnel plot*) mettant en lien la grandeur d'effet de chaque étude avec sa précision (nombre de participants).

Résultats (Jeanne Arminjon-Lupatelli)

À l'issue du processus de sélection et d'inclusion mené dans le cadre de cette revue systématique de littérature illustré par le diagramme de flux (Figure A), 14 articles ont pu être inclus. Sur les 208 publications sélectionnées au départ, 171 ont été exclues suite à l'analyse des titres et des résumés ou pour indisponibilité. Une lecture complète a été effectuée pour 37 articles, à la suite de laquelle 23 d'entre eux ont été exclus pour trois raisons principales : des données chiffrées ne pouvant pas être exploitées dans l'analyse statistique (n=11), l'absence de tâche FC « animaux » ou la passation de cette épreuve en 90 secondes au lieu de 60 secondes (n=6), ou un sujet trop éloigné de notre question (n=6).

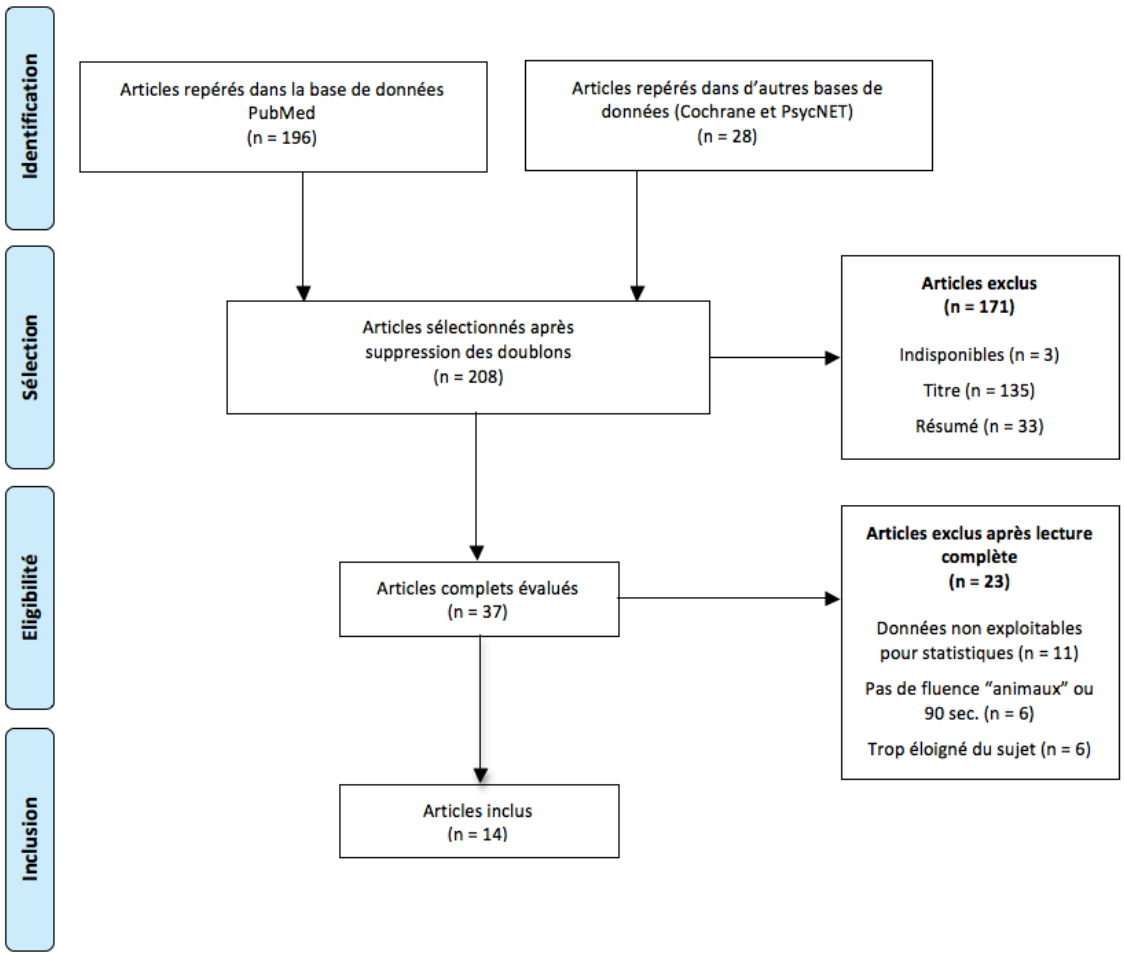

Figure A : Diagramme PRISMA (diagramme de flux ou *flow chart*).
D'après Moher, Liberati, Tetzlaff et Altman, 2009.

Tableau 3. Indicateurs de qualité des études incluses.
Qualité élevée : + ; incertaine : +/- ; faible : - . Classification HAS des études : CNR : Essai contrôlé non randomisé (grade B : présomption scientifique) ; ER : étude rétrospective (grade C : faible niveau de preuve). Critères MCI : Petersen (P) ; Winblad (W). Recrutement population contrôle : centre(s) mémoire (M). Âge : A2 (60-79 ans) ; A3 (80-95 ans). Education : E1 (0-8 ans d'études) ; E2 (9-12 ans d'études) ; E3 (13-21 ans d'études). NP : non précisé.

	Alegret	**Clark**	**Garcia-Herranz**	**Hall**	**Kim**	**Mc Donnell**	**Mirandez**	**Price**	**Rinehardt**	**Teng**	**Tröger**	**Vaughan**	**Wajman**	**Yeung**
Année parution	2018	2016	2019	2011	2014	2019	2017	2012	2014	2013	2019	2018	2018	2016
Revue	Journal of Alzheimer's disease	Alzheimer's and dementia	Aging, neuropsychology and cognition	Aging, neuropsychology and cognition	Psychiatry Investigation	International Psychogeriatrics	International Psychogeriatrics	Neuropsychology	Dementia and Geriatric cognitive disorders	Archives of clinical neuropsychology	Neuropsychologia	The American Geriatric Society	Applied neuropsychology adult	Frontiers in aging neurosciences
Impact factor	3.517	**14.423**	1.919	1.919	1.333	2.478	2.478	2.477	2.260	2.226	2.872	4.113	NP	3.633
Pays	Espagne	Etats-Unis	Espagne	Etats-Unis	Corée	Etats-Unis	Brésil	Australie	Etats-Unis	Etats-Unis	Allemagne France Suisse	Irlande	Brésil	Chine Australie
Type d'étude HAS	CNR	CNR	CNR	CNR	CNR	CNR	CNR	CNR	ER	CNR	CNR	CNR	CNR	CNR
Participants	**1820**	158	**287**	**488**	**1515**	**232**	67	66	136	**399**	95	**610**	**236**	52
Critères MCI	+ (P)	+ (P)	+ (P)	+ (W)	+ (W)	+ (P)	+ (P)	+ (P)	+ (P)	+ (P)	+ (P)	+ (P)	+ (P)	+ (P)
Sous-types MCI	non	oui	oui	non	oui	non	non	non	oui	oui	non	oui	oui	non
Recrutement pop. contrôle	+ (M)	+/-	+/-	+/-	+/-	+ (M)	+/-	+/-	+	+/-	+ (M)	+ (M)	+ (M)	+/-
Données démog.	+	+	+	+	+	+	+	+	+	+	+	+/-	+	+
Homogénéité âge	+ (A2)	+ (A2)	+ (A2)	+/- (A2, A3)	+ (A2)	+ (A2)	+ (A2)	+ (A2)	+ (A2)	+ (A2)	+ (A2)	+ (A2)	+ (A2)	+ (A2)
Homogénéité éducation	+/- (E1, E2)	+ (E3)	- (E1, E2, E3)	+/- (E2, E3)	+ (E1)	+ (E3)	+/- (E2, E3)	+/- (E2, E3)	+ (E3)	+ (E3)	+ (E2)	-	+/- (E1, E2)	+/- (E1, E2)
Démarche expé. explicite	+	+	+	+	+	+/-	+	+	+	+/-	+	+	+	+
Recueil résultats explicite	-	+/-	-	-	+	-	+	-	+	+/-	+/-	+	+	+/-
Présentation des résultats	+/-	+	+	+/-	+	+	+	+	+/-	+/-	+	+	-	+

Données démographiques

Le tableau 4 présente la synthèse des données démographiques pour l'ensemble des articles inclus. Dans le groupe MCI, la taille moyenne des effectifs est de 175 sujets avec cependant une variation importante entre les études, de 26 à 885 participants. Dans le groupe contrôle, la taille moyenne des effectifs est de 159 sujets avec également une variation importante de 20 à 655 participants. Au sein de chaque étude, le nombre de participants est réparti de manière assez homogène entre les deux groupes sauf pour quatre d'entre elles dont les effectifs par groupe vont du simple au double (Clark *et al.*, 2016 ; Hall, Harvey, Vo et O'Bryant, 2011 ; Rinehardt *et al.*, 2014 ; Tröger *et al.*, 2019).

L'âge moyen total est de 71,89 ans avec peu d'écart inter-groupes (72,54 ans pour les MCI et 71,18 ans pour les contrôles). Les deux groupes confondus, on relève un âge minimum de 63,3 ans (groupe contrôle dans Alegret *et al.*, 2018) et un âge maximum de 77,61 ans (groupe MCI dans Price *et al.*, 2012).

Concernant le sexe, les études comprennent entre 23,08% et 57% d'hommes dans le groupe MCI et entre 25% et 56% d'hommes dans le groupe contrôle. Le pourcentage d'hommes est en moyenne presque le même dans les deux groupes (40,15% pour les sujets MCI et 38,84% pour les contrôles).

Le nombre d'années d'éducation moyen est de 12,23 ans pour le groupe MCI (écart-type de 2,87) et de 13,08 ans pour le groupe contrôle (écart-type de 3,13). Le nombre d'années d'études minimum est équivalent pour le groupe MCI (7,94 ans) et le groupe contrôle (7,9 ans). Les deux groupes ont également un nombre d'années d'études maximum proche (16,06 ans pour les MCI et 17,6 ans pour les NC) avec un écart d'environ 10 années d'études entre minimum et maximum.

Tableau 4. Synthèse des données démographiques des études incluses

	Moy. MCI	**Min. MCI**	**Max. MCI**	**E-T MCI**	**Moy. NC**	**Min. NC**	**Max. NC**	**E-T NC**	**Moy. TOT.**	**Min. TOT.**	**Max. TOT.**	**E-T TOT.**
Taille des échantillons	175	26	885	229,17	159	20	655	206,51	166,96	20	885	214,20
Âge (ans)	72,54	64,77	77,61	3,44	71,18	63,3	77,3	3,72	71,89	63,3	77,61	3,57
Sexe (% hommes)	40,15	23,08	57	12,47	38,84	25	56	11,82	39,49	23,08	57	11,94
Éducation (années)	12,23	7,94	16,06	2,87	13,08	7,9	17,6	3,13	12,65	7,9	17,6	2,97

Analyse des données extraites pour le test des FC

Analyse de la grandeur d'effet

Une analyse statistique a été menée afin de confirmer ou d'infirmer notre première hypothèse : le score à la tâche FC « animaux » est significativement différent entre sujets MCI et sujets contrôles (NC).

Une mesure de l'effet (z-score) a été réalisée, représentée par un *forest plot* (Figure B). Celui-ci montre pour les 14 études incluses un effet significatif de la tâche FC « animaux » pour différencier les sujets MCI et NC car l'ensemble des lignes horizontales représentant l'intervalle de confiance pour chaque étude ne coupe pas la ligne verticale à 0 (« *line of no effect* »). L'estimation de l'effet de chaque étude est représentée par un rectangle dont l'aire est proportionnelle au poids de l'étude dans l'analyse statistique. Notons qu'aucune des études retenues ne chevauche la ligne verticale, ce qui révèle un consensus important. Le losange représente l'effet commun de toutes les études combinées. Comme il ne chevauche pas la ligne verticale et s'en éloigne même assez nettement, on peut considérer que le test a un effet significatif. L'hypothèse nulle d'une différence non significative entre m_1 (moyenne FC du groupe MCI) et m_2 (moyenne FC du groupe NC) est donc rejetée avec $p<0,00001$. Nous pouvons conclure que les moyennes aux FC « animaux » diffèrent entre les groupes MCI et contrôles de manière significative.

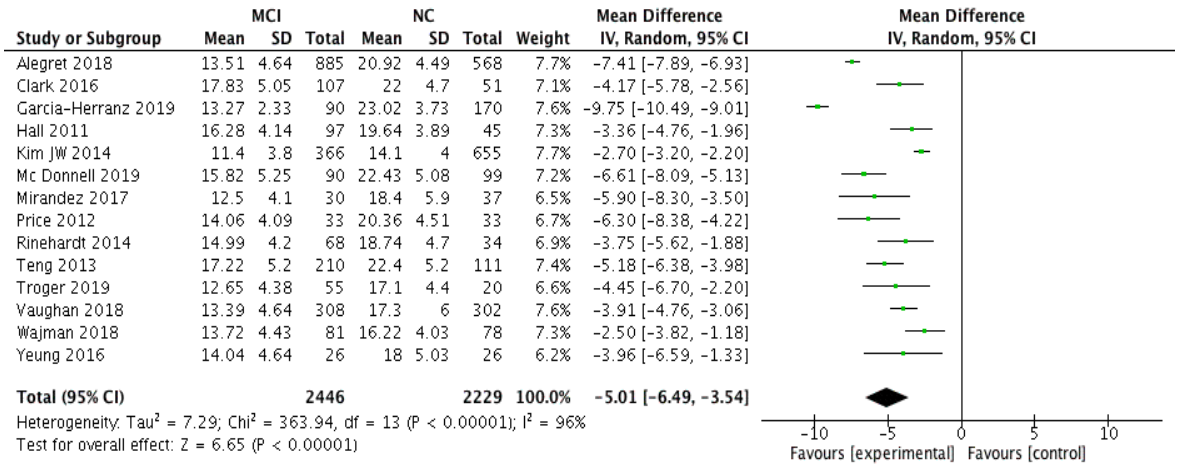

Figure B : *Forest plot* des scores à l'épreuve des FC pour les 14 études incluses (RevMan5).

Mise en évidence d'une hétérogénéité

Cependant, cet effet commun significatif est à nuancer par l'hétérogénéité des 14 études incluses, avec un $I^2 = 96\%$, soit une hétérogénéité importante (Zaugg *et al.,* 2014 ; Moulis et Sommet, 2014). En raison de cette valeur importante de I^2, la grandeur d'effet mentionnée ci-dessus a été calculée à l'aide d'un modèle d'analyse à effets aléatoires, préconisé afin de prendre en compte une telle hétérogénéité.

Comparaison des scores FC « animaux » et MMSE

Une analyse de la grandeur d'effet pour le test MMSE a été menée avec pour objectif de répondre à notre seconde hypothèse : la tâche des FC serait plus discriminante que le MMSE dans le diagnostic du MCI. Elle ne prend en compte que les 12 études pour lesquelles des scores au MMSE étaient disponibles parmi les 14 incluses (Figure C). Nous relevons une nouvelle fois une hétérogénéité importante avec I^2 à 76%, prise en compte par une analyse à effets aléatoires. Comme pour les FC, les moyennes au MMSE diffèrent entre les groupes MCI et NC de manière significative avec $p<0.00001$.

Study or Subgroup	MCI Mean	MCI SD	MCI Total	NC Mean	NC SD	NC Total	Weight	Mean Difference IV, Random, 95% CI
Alegret 2018	27.42	1.91	885	29.41	0.92	568	12.1%	-1.99 [-2.14, -1.84]
Clark 2016	27.26	2.01	107	28.9	1.2	51	8.7%	-1.64 [-2.14, -1.14]
Hall 2011	27.03	2.87	97	29.39	0.77	45	7.6%	-2.36 [-2.97, -1.75]
Kim JW 2014	23	4	366	25.3	3.2	655	9.0%	-2.30 [-2.78, -1.82]
Mc Donnell 2019	27.96	1.69	90	29.03	1.19	99	9.6%	-1.07 [-1.49, -0.65]
Mirandez 2017	27.8	1.7	30	29.2	1	37	6.9%	-1.40 [-2.09, -0.71]
Price 2012	27.4	1.4	33	29	0.9	33	8.1%	-1.60 [-2.17, -1.03]
Rinehardt 2014	27.15	2.15	68	28.88	1.2	34	7.3%	-1.73 [-2.38, -1.08]
Teng 2013	27.3	2.09	210	29.1	1.2	111	10.2%	-1.80 [-2.16, -1.44]
Troger 2019	25.9	2.7	55	28.3	1.7	20	4.4%	-2.40 [-3.43, -1.37]
Vaughan 2018	24.6	2.8	308	27.3	2.6	302	9.5%	-2.70 [-3.13, -2.27]
Wajman 2018	26.6	2.69	81	27.62	1.97	78	6.5%	-1.02 [-1.75, -0.29]
Total (95% CI)			**2330**			**2033**	**100.0%**	**-1.84 [-2.11, -1.57]**

Heterogeneity: $Tau^2 = 0.15$; $Chi^2 = 45.70$, df = 11 ($P < 0.00001$); $I^2 = 76\%$
Test for overall effect: Z = 13.34 ($P < 0.00001$)

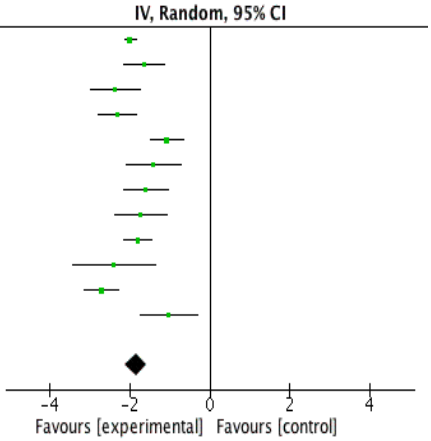

Figure C : *Forest plot* des scores à l'épreuve du MMSE pour 12 études (RevMan5).

Les écarts de moyennes entre les groupes contrôles et MCI sont donc significatifs que ce soit pour les FC ou le MMSE.

Pour comparer ces deux tests, il est possible de calculer un « d de Cohen ». Cet indicateur suggère qu'au sein des 12 études retenues pour cette comparaison, l'écart des moyennes entre les deux groupes est très légèrement plus important pour les FC « animaux » que pour le MMSE (Tableau 5). Cependant, l'écart est minime et peu significatif. Nous relevons que les deux tailles d'effet sont fortes.

Tableau 5. Comparaison des FC au MMSE par le calcul du « d de Cohen ».

12 études MMSE + FC	Groupe 1_MCI	Groupe 2_NC
Nombre de sujets	2330	2033
FC_moyenne	14.45	19.13
FC_SD	4.49	4.74
d Cohen_FC	**1.01 => taille d'effet : forte** (>0.8)	
MMSE_moyenne	26.62	28.45
MMSE_SD	2.33	1.49
d Cohen_MMSE	**0.94 => taille d'effet : forte** (>0.8)	

Exploration de l'hétérogénéité

Recrutement et évaluation des sujets MCI

L'hétérogénéité importante présente dans les deux parties de notre analyse pourrait affaiblir nos conclusions, malgré une taille d'effet significative. Nous nous sommes attachées à rechercher les causes de cette variabilité à travers l'analyse qualitative des articles inclus. Ainsi, l'hétérogénéité inhérente au concept de MCI, composé de sous-types en lien avec une diversité étiologique, fait partie de ces causes probables. Cette diversité clinique et étiologique se retrouve dans les 14 études retenues (Tableau 6).

Tableau 6. Synthèse par étude des types de MCI et de démences.
aMCI-sd/md : *amnestic MCI single/multiple domain(s)*. naMCI-sd/md : *non amnestic MCI single/multiple domain*(s). MA : maladie d'Alzheimer. DFT : démence fronto-temporale. DCL : démence à corps de Lewy. DV : démence vasculaire. DP : dépression. MP : maladie de Parkinson.

	Sous-types MCI		Présence d'un groupe avec démence	
Alegret	1	MCI (non spécifié)	1	MA légère
Clark	4	aMCI-sd, aMCI-md, naMCI-sd, naMCI-md	0	-
Garcia-Herranz	1	MCI (non spécifié)	1	MA
Hall	1	MCI (non spécifié)	2	MA, DV
Kim	2	aMCI, naMCI	5	MA, DV, DCL, MP, DFT
Mc Donnell	1	aMCI	1	MA
Mirandez	1	MCI (non spécifié)	0	-
Price	1	MCI (non spécifié)	0	-
Rinehardt	2	aMCI, naMCI	1	MA
Teng	2	aMCI, naMCI	1	MA
Tröger	1	MCI (non spécifié)	1	MA
Vaughan	4	aMCI-sd, aMCI-md, naMCI-sd, naMCI-md	0	-
Wajman	2	aMCI-sd, aMCI-md	3	MA, DCL, DFT
Yeung	1	MCI (non spécifié)	0	-

Outre cette variabilité clinique et étiologique, on relève une variabilité dans le mode d'évaluation des sujets MCI (tests et seuils pathologiques choisis, évaluateurs, double

aveugle ou non) ou de leur provenance : études épidémiologiques, centre(s) mémoire (un ou plusieurs avec chacun ses spécificités).

Analyse des données démographiques

La variabilité de la taille des échantillons de chaque étude (de 52 à 1820 participants) a un impact sur la précision des résultats et sur le poids de chaque étude dans l'effet commun. Certaines études ont des effectifs très variables selon les groupes (jusqu'à deux fois plus de sujets MCI que de contrôles dans certaines d'entre elles, l'inverse pour d'autres).

Les études incluses affichent une certaine hétérogénéité en terme de sexe mais cette variable n'influerait que modérément sur les scores aux FC (Auriacombe *et al.*, 2001).

Les moyennes d'âge des deux groupes apparaissent quant à elles assez homogènes et se répartissent sur seulement deux des quatre tranches d'âge définies dans l'étude d'Auriacombe *et al.* (2001) : ≤ 74 ans (moyenne FC : 31.4 ; écart-type : 7.6) et 75-79 ans (moyenne FC : 29.2 ; écart-type : 6.9). Les moyennes et écarts-types de ces deux tranches d'âge restent proches (ce qui n'est pas le cas au-delà de 80 ans). En outre, la majorité des participants des 14 études se situent dans la première tranche avec un âge moyen inférieur ou égal à 74 ans (ce qui concerne 10 études pour le groupe contrôle et 9 pour le groupe MCI).

Comme nous l'avons relevé dans le tableau des indicateurs de qualité des études incluses (Tableau 3) et à travers la présentation des données démographiques (Tableau 4), le niveau d'éducation peut différer de manière significative entre les études retenues et parfois au sein d'une même étude. Par ailleurs, les articles étant issus de pays différents avec leur propre système éducatif, le nombre d'années d'études mentionné ne reflète pas forcément la même réalité selon les pays. Le niveau d'éducation constitue donc une source d'hétérogénéité probable dans cette revue systématique.

Les articles inclus n'ont pas fourni de données concernant le métier exercé (ou exercé par le passé) par les participants, il n'a donc pas été possible d'exploiter cette variable.

Biais de publication

La recherche d'un éventuel biais de publication est une étape nécessaire dans le cadre d'une revue systématique de littérature. Ce type de biais est notamment dû à une certaine sélectivité dans la publication des articles pour ne publier que ceux dont les résultats sont significatifs ou favorables à l'hypothèse de départ (Maison, 2010). Ce biais est traditionnellement représenté sous forme d'un graphique en entonnoir ou *funnel plot*, avec en ordonnée, la précision de l'étude (taille de l'échantillon), et en abscisse, l'effet de

l'intervention. S'il n'a pas une forme symétrique et pyramidale, cela suggère un probable biais de publication (Zaugg *et al.*, 2014).

Le *funnel plot* des 14 articles inclus est représenté ci-dessous (Figure D). Il montre un potentiel biais de publication car certaines études sortent de l'entonnoir bien que globalement distribuées de part et d'autre de l'effet commun.

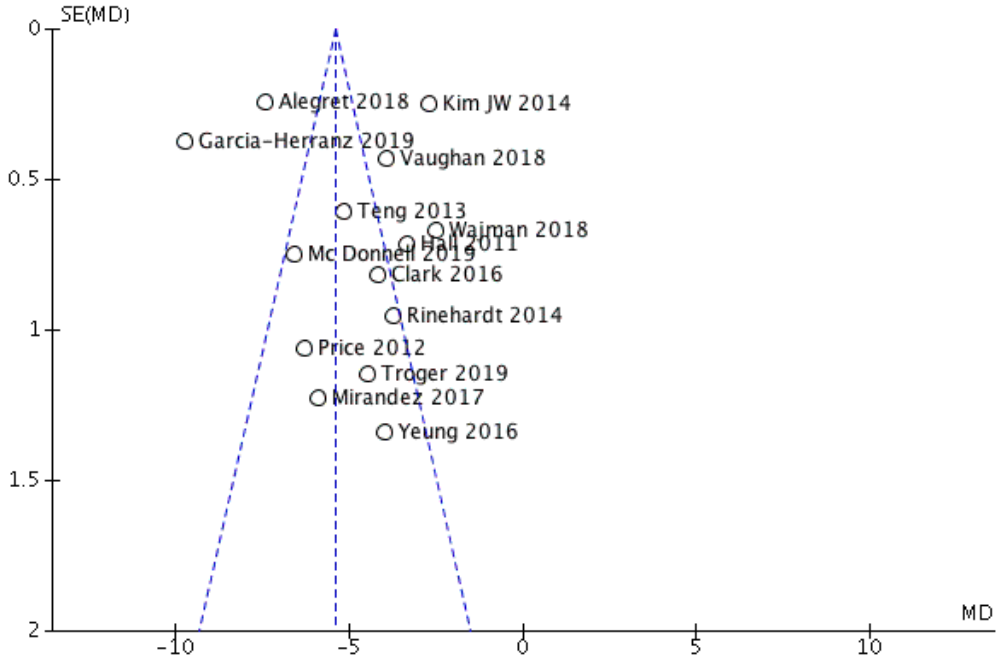

Figure D : Biais de publication représenté sous forme de *funnel plot* (RevMan 5).

Discussion et conclusion (Jeanne Arminjon-Lupatelli)

Cette revue systématique de littérature visait à déterminer si l'épreuve de fluence catégorielle était un outil pertinent pour différencier les sujets contrôles des MCI de manière précoce, en elle-même et par comparaison au MMSE.

FC et/ou MMSE pour dépister un MCI ?

Dans tous les articles inclus, les sujets MCI obtiennent des résultats significativement différents des contrôles à la tâche FC « animaux », confirmant la pertinence de celle-ci dans le diagnostic du MCI. Le calcul d'une grandeur d'effet à partir des données extraites des 14 études qui représentent au total 4675 participants (2446 MCI et 2229 contrôles) nous a permis de conclure à une différence significative entre les moyennes aux FC « animaux » des groupes MCI et contrôles. Nous avons pu confirmer notre hypothèse principale, à savoir que l'épreuve de FC « animaux » est un instrument de *screening* adapté pour dépister les patients MCI.

Concernant le MMSE, l'analyse a concerné 12 études proposant ce test parmi les 14 incluses soit 4363 participants (2330 MCI et 2033 contrôles). Là encore, des écarts de moyennes significatifs ont été mis en évidence entre les deux groupes. Par conséquent, il est possible de conclure que les FC comme le MMSE sont des outils adaptés pour dépister un trouble neurocognitif au stade précoce. En revanche, nous n'avons pas pu montrer de différence significative entre les FC et le MMSE pour ce dépistage, les « d de Cohen » ayant révélé pour chacun des deux tests une taille d'effet forte. Cette revue systématique s'appuyant sur des données résumées, sans accès aux données individuelles, il n'a pas été possible de mettre en exergue un éventuel effet combiné de la passation des deux tests.

Toutefois, plusieurs études incluses nuancent et questionnent l'intérêt du MMSE pour mettre en évidence le MCI. S'il reste l'outil de mesure de l'efficience cognitive globale le plus répandu à travers le monde, il ne serait pas le test le plus adéquat pour dépister les troubles neurocognitifs au stade précoce. En effet, il évalue peu les fonctions exécutives et n'explore pas assez finement les mémoires épisodique et sémantique. En outre, il ne permet pas toujours de dépister les sujets MCI qui peuvent obtenir un score dans la norme surtout s'ils ont un haut niveau d'éducation (Hwang *et al.*, 2019).

L'association d'une épreuve de FC au MMSE pourrait permettre à ce dernier d'être plus performant dans la détection précoce d'un trouble neurocognitif léger. En effet, l'ajout d'une épreuve de FC « animaux » accroît la sensibilité du MMSE (de 70.1 à 81.2%) et sa spécificité (de 71.7 à 77.8%) dans le diagnostic d'un aMCI (variante amnésique du MCI) ou d'une MA préclinique (Mc Donnell *et al.*, 2019). La capacité de détection du MCI, quel

que soit son sous-type, serait donc accrue par la combinaison de ces deux tests qui est significativement meilleure que celle du MMSE seul (Kim *et al.*, 2014).

En outre, le test FC « animaux » présente un avantage d'ordre pratique : avec un temps de passation inférieur à 5 minutes et une grande facilité d'administration, cette épreuve pourrait se généraliser en tant que test de *screening* à disposition par exemple des médecins traitants, des neurologues, des orthophonistes et des neuropsychologues.

Fluences verbales et trouble neurocognitif léger

Plusieurs études incluses ont comparé les scores aux fluences phonémiques avec les scores aux fluences catégorielles obtenus par deux ou trois groupes de sujets : contrôles, MCI et atteints de démence. Dans la population générale, y compris vieillissante, la fluence catégorielle est généralement mieux réussie que la fluence phonémique. Cependant, ce ratio est inversé pour les patients atteints par la maladie d'Alzheimer (Vaughan, Coen, Kenny et Lawlor, 2018). Les deux types de fluences font appel aux fonctions exécutives largement dépendantes du lobe frontal, mais la fluence catégorielle sollicite aussi la mémoire sémantique qui serait perturbée précocement en cas de lésions temporales, comme c'est le cas dans la maladie d'Alzheimer (Rinehardt *et al.,* 2014 ; Teng *et al.*, 2013 ; Vaughan *et al.*, 2018).

Une majorité d'études ayant analysé l'écart de performances entre FC et FP chez le sujet MCI, et plus particulièrement aMCI, ont montré un *pattern* similaire à ce qui est retrouvé dans la MA : de moins bonnes performances en FC qu'en FP, même si elles restent meilleures que celles des sujets avec une maladie d'Alzheimer. Ces résultats suggèrent que ce ratio FP/FC pourrait permettre de différencier les sujets MCI des sujets sains et prédire le risque d'évolution vers une démence de type MA (García-Herranz, Díaz-Mardomingo, Venero et Peraita, 2019 ; Teng *et al.,* 2013 ; Vaughan *et al.*, 2018). Il n'a pas été possible d'analyser statistiquement ce ratio dans notre revue systématique car les études incluses ne comportaient pas toutes des scores en FP ou, si elles en comportaient, les tâches de FP étaient trop différentes pour pouvoir être comparées.

La FC « animaux », la plus largement utilisée, est celle qui a montré la plus forte sensibilité pour différencier les sujets contrôles, MCI et MA (García-Herranz *et al*., 2019). Alors que les sujets MCI présentent des performances significativement inférieures aux contrôles pour tous les types de fluences, la FC « animaux » montre la plus forte corrélation avec la concentration de protéine tau dans le liquide céphalo-rachidien, l'un des principaux marqueurs de la MA (Mirandez, Aprahamian, Talib, Forlenza et Radanovic, 2017).

L'analyse qualitative de ce type de FC révèle que les patients MCI génèrent moins de mots, moins de nouvelles sous-catégories et des sous-catégories plus proches sémantiquement les unes des autres par rapport aux sujets contrôles et produisent aussi de plus petits *clusters* (regroupements sémantiques), ce qui témoigne d'une dégradation précoce des stratégies d'exploration et de l'accès au système sémantique (Price *et al.*, 2012 ; Tröger *et al.*, 2019 ; Wajman, Cecchini, Bertolucci et Mansur, 2018 ; Yeung *et al.*, 2016).

Limites

La principale limite de cette revue systématique concerne l'hétérogénéité importante des articles inclus. Néanmoins, « l'hétérogénéité est inhérente à la méta-analyse et ne doit pas être considérée comme un problème devant être éliminé, mais au contraire comme un paramètre très informatif pour l'extrapolation des données » (Zaugg *et al.*, 2014). Au-delà du traitement statistique, les sources d'hétérogénéité sont elles-mêmes informatives et il appartient au travail de revue systématique de les rechercher au travers de l'analyse qualitative des études incluses (Moulis et Sommet, 2014).

Une exploration fine des sources de variabilité étude par étude nous a permis de relever les principaux facteurs d'hétérogénéité : le recrutement des patients MCI (définition du MCI et éventuels sous-types, lieu et mode de recrutement) ; le recrutement des sujets contrôles (concept de vieillissement normal, lieu et mode de recrutement) ; la présence ou non d'une composante pronostique et d'un troisième groupe avec démence(s) constituée(s) ; les facteurs démographiques avec en premier lieu le niveau d'éducation.

En outre, ne disposant pas des scores individuels des participants, l'analyse statistique a dû être réalisée à partir des moyennes, ce qui engendre des résultats moins précis que si elle avait été réalisée à partir des scores individuels. Cependant, les méta-analyses sur données résumées sont plus courantes que les méta-analyses sur données individuelles qui nécessitent d'obtenir les bases de données de toutes les études incluses et sont donc plus difficiles à mettre en œuvre (Maison, 2010 ; Moulis et Sommet, 2014).

Enfin, nous avons limité nos recherches aux travaux publiés et référencés par trois grandes bases de données (MEDLINE/PubMed, Cochrane Library et APA PsycINFO/PsycNET). Il aurait été intéressant d'explorer la littérature « grise » pour limiter le risque de biais de publication mis en évidence par le *funnel plot* des 14 études incluses.

Perspectives

À l'avenir, il serait possible d'explorer plus finement les épreuves de FC en analysant les nombreux résultats présents dans la littérature en termes de *clustering* (regroupements sémantiques de mots) et de *switching* (passage d'un *cluster* à un autre). Il serait aussi pertinent de restreindre les critères d'inclusion des études en fonction du type de démence, du sous-type de MCI et/ou du niveau d'éducation des sujets afin de réduire les sources d'hétérogénéité. Il pourrait également être intéressant d'étudier la performance diagnostique de cette tâche associée à d'autres examens non invasifs et peu coûteux (par exemple la NIRS, *near-infrared spectroscopy*) ou à d'autres indicateurs neuropsychologiques (la mémoire de travail notamment). Examiner sa valeur pronostique au travers d'études longitudinales pour prédire l'évolution vers une démence serait un autre axe à développer.

Cette revue systématique de la littérature incluant 14 études menées ces dix dernières années a confirmé la pertinence de la tâche de fluence catégorielle « animaux » dans le cadre du diagnostic précoce du MCI, qu'il s'agisse d'un test de *screening* réalisé en première intention ou de l'intégrer à la batterie des tests neurocognitifs proposés lors d'une évaluation plus complète. Notre étude, qui s'appuie sur une grandeur d'effet impliquant 4675 participants, confirme que cette épreuve rapide et facile à administrer mériterait d'être systématisée dans le dépistage précoce des troubles neurocognitifs.

Bibliographie (APA, 6e édition)

Alegret, M., Peretó, M., Pérez, A., Valero, S., Espinosa, A., Ortega, G., …Boada, M. (2018). The role of verb fluency in the detection of early cognitive impairment in Alzheimer's disease. *Journal of Alzheimer's Disease*, *62*(2), 611-619.

Auriacombe, S., Fabrigoule, C., Lafont, S., Jacqmin-Gadda, H. et Dartigues, J.-F. (2001). Letter and category fluency in normal elderly participants: A population-based study. *Aging, Neuropsychology, and Cognition, 8*(2), 98-108.

Bettany-Saltikov, J. (2010a). Learning how to undertake a systematic review: Part 1. *Nursing Standard*, *24*(50), 47-55.

Bettany-Saltikov, J. (2010b). Learning how to undertake a systematic review: Part 2. *Nursing Standard*, *24*(51), 47-56.

Clark, D. G., McLaughlin, P. M., Woo, E., Hwang, K., Hurtz, S., Ramirez, L., …Apostolova, L. G. (2016). Novel verbal fluency scores and structural brain imaging for prediction of cognitive outcome in mild cognitive impairment. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *2*(1), 113-122.

Ebbels, S. H. (2017). Intervention research: Appraising study designs, interpreting findings and creating research in clinical practice. *International Journal of Speech-Language Pathology, 19*(3), 218-231.

García-Herranz, S., Díaz-Mardomingo, M. C., Venero, C. et Peraita, H. (2019). Accuracy of verbal fluency tests in the discrimination of mild cognitive impairment and probable Alzheimer's disease in older Spanish monolingual individuals. *Aging, Neuropsychology, and Cognition*, 1-15.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C. et Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children, 71*(2), 149-164.

Hall, J. R., Harvey, M., Vo, H. T. et O'Bryant, S. E. (2011). Performance on a measure of category fluency in cognitively impaired elderly. *Aging, Neuropsychology, and Cognition*, *18*(3), 353-361.

Hwang, A. B., Boes, S., Nyffeler, T. et Schuepfer, G. (2019). Validity of screening instruments for the detection of dementia and mild cognitive impairment in hospital inpatients: A systematic review of diagnostic accuracy studies. *PLOS ONE*, *14*(7), e0219569.

Kim, J. E., Park, S.-H., Hong, Y. J., Hwang, J., Han, N. E., Lee, S.-M., ...Lee, J.-H. (2019). Qualitative comparison of semantic memory impairment in patients with amnestic mild cognitive impairment based on β-amyloid status. *Journal of Clinical Neurology*, *15*(1), 27-37.

Kim, J. W., Lee, D. Y., Seo, E. H., Sohn, B. K., Choe, Y. M., Kim, S. G., …Woo, J. I. (2014). Improvement of screening accuracy of Mini-Mental State Examination for mild cognitive impairment and non-Alzheimer's disease dementia by supplementation of verbal fluency performance. *Psychiatry Investigation, 11*(1), 44.

Lehrner, J., Coutinho, G., Mattos, P., Moser, D., Pflüger, M., Gleiss, A., …Stögmann, E. (2017). Semantic memory and depressive symptoms in patients with subjective cognitive decline, mild cognitive impairment, and Alzheimer's disease. *International Psychogeriatrics, 29*(7), 1123-1135.

Maison, P. (2010). La méta-analyse sur données résumées. *Recherche en soins infirmiers, 101*(2), 18-24.

McDonnell, M., Dill, L., Panos, S., Amano, S., Brown, W., Giurgius, S., ...Miller, K. (2019). Verbal fluency as a screening tool for mild cognitive impairment. *International Psychogeriatrics*, 1-8.

Mirandez, R. M., Aprahamian, I., Talib, L. L., Forlenza, O. V. et Radanovic, M. (2017). Multiple category verbal fluency in mild cognitive impairment and correlation with CSF biomarkers for Alzheimer's disease. *International Psychogeriatrics*, *29*(6), 949-958.

Moher, D., Liberati, A., Tetzlaff, J. et Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of internal medicine*, *151*(4), 264-269.

Mortamais, M., Ash, J. A., Harrison, J., Kaye, J., Kramer, J., Randolph, C., ...Ritchie, K. (2017). Detecting cognitive changes in preclinical Alzheimer's disease: A review of its feasibility. *Alzheimer's & Dementia, 13*(4), 468-492.

Moulis, G. et Sommet, A. (2014). Comment lire une méta-analyse ? *La Revue de Médecine Interne*, *35*(4), 250-258.

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*(3), 183-194.

Price, S. E., Kinsella, G. J., Ong, B., Storey, E., Mullaly, E., Phillips, M., …Perre, D. (2012). Semantic verbal fluency strategies in amnestic mild cognitive impairment. *Neuropsychology*, *26*(4), 490-497.

Rinehardt, E., Eichstaedt, K., Schinka, J. A., Loewenstein, D. A., Mattingly, M., Fils, J., …Schoenberg, M. R. (2014). Verbal fluency patterns in mild cognitive impairment and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, *38*(1-2), 1-9.

Tangalos, E. G. et Petersen, R. C. (2018). Mild cognitive impairment in geriatrics. *Clinics in Geriatric Medicine, 34*(4), 563–589.

Teng, E., Leone-Friedman, J., Lee, G. J., Woo, S., Apostolova, L. G., Harrell, S., …Lu, P. H. (2013). Similar verbal fluency patterns in amnestic mild cognitive impairment and Alzheimer's disease. *Archives of Clinical Neuropsychology*, *28*(5), 400-410.

Tröger, J., Linz, N., König, A., Robert, P., Alexandersson, J., Peter, J. et Kray, J. (2019). Exploitation vs. exploration – computational temporal and semantic analysis explains semantic verbal fluency impairment in Alzheimer's disease. *Neuropsychologia, 131*, 53-61.

Venneri, A., Jahn-Carta, C., De Marco, M., Quaranta, D. et Marra, C. (2018). Diagnostic and prognostic role of semantic processing in preclinical Alzheimer's disease. *Biomarkers in medicine, 12*(6), 637-651.

Vaughan, R. M., Coen, R. F., Kenny, R. et Lawlor, B. A. (2018). Semantic and phonemic verbal fluency discrepancy in mild cognitive impairment: Potential predictor of progression to Alzheimer's disease. *Journal of the American Geriatrics Society, 66*(4), 755-759.

Wajman, J. R., Cecchini, M. A., Bertolucci, P. H. F. et Mansur, L. L. (2018). Quanti-qualitative components of the semantic verbal fluency test in cognitively healthy controls, mild cognitive impairment, and dementia subtypes. *Applied Neuropsychology: Adult*, *26*(6), 533-542.

Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L.-O., …Petersen, R. C. (2004). Mild cognitive impairment – beyond controversies, towards a consensus: Report of the international working group on mild cognitive impairment. *Journal of Internal Medicine, 256*(3), 240-246.

Yeung, M. K., Sze, S. L., Woo, J., Kwok, T., Shum, D. H. K., Yu, R. et Chan, A. S. (2016). Altered frontal lateralization underlies the category fluency deficits in older adults with mild cognitive impairment: A near-infrared spectroscopy study. *Frontiers in Aging Neuroscience*, *8*.

Zaugg, V., Savoldelli, V., Sabatier, B. et Durieux, P. (2014). Améliorer les pratiques et l'organisation des soins : méthodologie des revues systématiques. *Santé Publique*, *26*(5), 655-667.

Annexe A – Synthèse des caractéristiques des articles inclus.
MCI : *mild cognitive impairment*. NC : sujets contrôles. NB : nombre de sujets. Âge : âge moyen. NE : niveau moyen d'éducation en années d'études. NP : non précisé. FV : fluence verbale. FP : fluence phonémique. FC : fluence catégorielle. MA : maladie d'Alzheimer. aMCI : MCI variante amnésique. Non-aMCI : MCI variante non-amnésique.

Auteurs (date de publication)	**Titre article**	**Population MCI**			**Population NC**			**Résultats/Conclusions**
		NB	Âge	NE	NB	Âge	NE	
Alegret, M., Pereto, M., Perez, A., Valero, S., Espinosa, A., Ortega, G., ...Boada, M. (2018)	The role of verb fluency in the detection of early cognitive impairment in Alzheimer's disease	885	71,7	9,3	568	63,3	11,9	Les fluences de verbes pourraient être un bon marqueur linguistique pour le diagnostic différentiel entre simple plainte mnésique chez le sujet âgé et MCI.
Clark, D. G., McLaughlin, P. M., Woo, E., Hwang, K., Hurtz, S., Ramirez, L., ...Apostolova, L. G. (2016)	Novel verbal fluency scores and structural brain imaging for prediction of cognitive outcome in mild cognitive impairment	107	71,25	16	51	68,9	17,6	L'analyse détaillée des FV peut permettre de développer une méthode de détection rapide, peu chère et non invasive des changements cognitifs précoces de la MA. Les mesures ayant trait à l'analyse lexicale des mots produits (fréquence lexicale, cohérence) apparaissent comme ayant le plus fort pouvoir prédictif du risque d'évolution vers une MA.
García-Herranz, S., Díaz-Mardomingo, M. C., Venero, C. et Peraita, H. (2019)	Accuracy of verbal fluency tests in the discrimination of mild cognitive impairment and probable Alzheimer's disease in older Spanish monolingual individuals	90	70,76	11,44	170	68,36	14,42	Les FV permettent de différencier les sujets NC et MCI. Les FC sont davantage déficitaires que les FP chez les sujets MCI. Parmi cinq tâches de FV, la FC « animaux » a montré la plus grande capacité à différencier les profils, y compris sujets MCI vs sujets avec une MA très légère. C'est aussi la FC « animaux » qui a montré la meilleure capacité à prédire l'évolution vers une MA.
Hall, J. R., Harvey, M., Vo, H. T. et O'Bryant, S. E. (2011)	Performance on a measure of category fluency in cognitively impaired elderly	97	76,86	14,56	45	73,74	15,05	Les écarts de scores des FC émergent à des périodes différentes de l'épreuve selon les groupes diagnostiques : - la première période (0-15 s.) permet de faire la différence entre les groupes « déments » (MA et démence vasculaire) et « non déments » (NC et MCI) ; - la deuxième période (15-30 s.) permet de faire la différence entre le vieillissement normal et le MCI.

Auteurs (date de publication)	Titre article	Population MCI			Population NC			Résultats/Conclusions
		NB	Âge	NE	NB	Âge	NE	
Kim, J. W., Lee, D. Y., Seo, E. H., Sohn, B. K., Choe, Y. M., Kim, S. G., ...Woo, J. I. (2014)	Improvement of screening accuracy of Mini-Mental State Examination for mild cognitive impairment and non-Alzheimer's disease dementia by supplementation of verbal fluency performance	366	71,1	8,3	655	69,5	7,9	Le MMSE et les FC « animaux » ont une capacité de détection significativement meilleure par rapport au MMSE seul pour : - détecter le MCI quel que soit son sous-type (vs NC) - détecter les démences d'autres types que la MA (DFT, démence vasculaire...) - pas de gain de sensibilité significatif pour détecter les patients MA
McDonnell, M., Dill, L., Panos, S., Amano, S., Brown, W., Giurgius, S., ...Miller, K. (2019)	Verbal fluency as a screening tool for mild cognitive impairment	90	64,77	16,06	99	63,48	16,3	Ajouter une tâche de FC animaux au MMSE augmente la sensibilité et la spécificité de celui-ci pour le diagnostic du MCI.
Mirandez, R. M., Aprahamian, I., Talib, L. L., Forlenza, O. V. et Radanovic, M. (2017)	Multiple category verbal fluency in mild cognitive impairment and correlation with CSF biomarkers for Alzheimer's disease	30	74,5	12	37	72,5	13,5	Il existe un lien entre hypoperfusion du lobe temporal, bas niveau d'éducation et performances altérées en fluences de verbes et entre hypoperfusion de la région temporo-occipitale, vieillissement normal et performances altérées en FC classiques. Biomarqueurs : niveaux de T-tau corrélés aux scores des FC « animaux ».
Price, S. E., Kinsella, G. J., Ong, B., Storey, E., Mullaly, E., Phillips, M., ...Perre, D. (2012)	Semantic verbal fluency strategies in amnestic mild cognitive impairment	33	77,61	12,15	33	75,52	13,46	Les sujets MCI montrent des difficultés à isoler les catégories sémantiques et on note une perte de liens associatifs au sein des catégories sémantiques (plus petits clusters, moins de sous-catégories que les NC). Ces résultats seraient à mettre en relation avec une dégradation précoce de la mémoire sémantique dans le cadre du aMCI.
Rinehardt, E., Eichstaedt, K., Schinka, J. A., Loewenstein, D. A., Mattingly, M., Fils, J., ...Schoenberg, M. R. (2014)	Verbal fluency patterns in mild cognitive impairment and Alzheimer's disease	68	74,45	13,85	34	75,21	13,74	Sujets NC et MCI : FC>FP (même si moins bonnes performances des MCI) alors que sujets MA : FC<FP. Les sujets MCI ont un ratio FC/FV plus proches des NC que des MA. Ce ratio ne permettrait donc pas de détecter la MA précoce chez des sujets MCI.

Auteurs (date de publication)	Titre article	Population MCI			Population NC			Résultats/Conclusions
		NB	Âge	NE	NB	Âge	NE	
Teng, E., Leone-Friedman, J., Lee, G. J., Woo, S., Apostolova, L. G., Harrell, S., ...Lu, P. H. (2013)	Similar verbal fluency patterns in amnestic mild cognitive impairment and Alzheimer's disease	210	73,05	16	111	70,3	17	Les sujets aMCI présentent un déficit plus important en FC qu'en FP comparé aux NC. Ce déficit supérieur en FC chez les aMCI est comparable à ce qui est retrouvé dans la MA stade léger mais de manière moins prononcée. Les non-aMCI, quant à eux, présentent un ratio FP/FC similaires à celui des NC. *Patterns* neuropathologiques sous-jacents différents entre aMCI et non-aMCI, avec en particulier la prédominance de la dégénérescence précoce du lobe temporal dans le aMCI.
Tröger, J., Linz, N., König, A., Robert, P., Alexandersson, J., Peter, J. et Kray, J. (2019)	Exploitation vs. exploration-computational temporal and semantic analysis explains semantic verbal fluency impairment in Alzheimer's disease	55	77	10,8	20	77,3	10,9	Les résultats obtenus par l'approche computationnelle modélisent les difficultés que les sujets rencontrent au sein des *clusters* ou entre eux en termes d'efficacité et d'efficience. Les processus de *clustering* et de *switching* se révèlent inefficaces et chronophages chez les patients MA, ce qui est en relation avec le déficit global des scores de FC chez ces patients. Un *pattern* similaire a été trouvé dans une moindre mesure chez les patients MCI (sans significativité montrée sur le plan statistique).
Vaughan, R. M., Coen, R. F., Kenny, R. et Lawlor, B. A. (2018)	Semantic and phonemic verbal fluency discrepancy in mild cognitive impairment: Potential predictor of progression to Alzheimer's disease	308	71,3	NP	302	71,6	NP	Les sujets MCI pour qui FP>FC ont plus de risque de progresser vers une MA que ceux pour qui FC>FP. L'écart entre les scores en FC et FP est un indicateur de la progression potentielle de sujets MCI vers une MA.
Wajman, J. R., Cecchini, M. A., Bertolucci, P. H. F. et Mansur, L. L. (2018)	Quanti-qualitative components of the semantic verbal fluency test in cognitively healthy controls, mild cognitive impairment, and dementia subtypes	81	72,11	10,6	78	70,29	8,63	Intérêt de l'analyse qualitative des FC, en complément de l'analyse quantitative. Le nombre de mots produits a montré la meilleure sensibilité (87,2%) tandis que la taille des *clusters* a montré la meilleure spécificité (94,9%) pour différencier les sujets aMCI des MA et pour différencier les MA des NC. La variable la plus efficace pour différencier NC et MCI serait la taille moyenne des *clusters*.
Yeung, M. K., Sze, S. L., Woo, J., Kwok, T., Shum, D. H. K., Yu, R. et Chan, A. S. (2016)	Altered frontal lateralization underlies the category fluency deficits in older adults with mild cognitive impairment: A near-infrared spectroscopy study	26	69,07	7,94	26	68,87	9,58	Activité du lobe préfrontal significativement latéralisée à gauche pour les NC pendant une tâche de FC alors qu'elle semble bilatérale chez les sujets MCI. Une moindre latéralisation de l'activité frontale serait donc à l'origine des déficits en FC chez les sujets MCI, en particulier ceux appartenant à la catégorie non-amnésique.

Apport des fluences catégorielles dans le diagnostic du *mild cognitive impairment :* revue systématique de littérature.

Résumé

Le concept de *mild cognitive impairment* (MCI) ou trouble neurocognitif léger se définit comme une phase de transition entre le vieillissement normal et l'entrée dans une pathologie neurodégénérative. Le diagnostic précoce des patients MCI est un enjeu majeur de santé publique afin d'anticiper une prise en charge adaptée.

À l'issue d'un processus rigoureux, cette revue systématique de littérature a retenu 14 études publiées ces dix dernières années impliquant 4675 participants (2446 MCI et 2229 sujets contrôles). Elle vise à déterminer si l'épreuve de fluence catégorielle (FC) est un outil pertinent pour le diagnostic précoce du MCI. L'analyse statistique réalisée à partir des données extraites a permis de conclure à une différence significative entre les moyennes aux FC « animaux » des groupes MCI et contrôles, confirmant la pertinence de l'épreuve de fluence catégorielle dans le cadre du diagnostic précoce du MCI.

Mots clés : trouble neurocognitif léger, fluences verbales, fluences catégorielles, évaluation, diagnostic

Abstract

Mild cognitive impairment (MCI) is defined as a transitional phase between normal aging and the onset of a neurodegenerative pathology. Early diagnosis of MCI patients is a critical public health issue in order to anticipate an appropriate treatment.

Through a rigorous selection process, fourteen articles published in the last ten years and involving 4675 participants (2446 MCI and 2229 controls) were included in this systematic review. The aim was to determine whether the semantic fluency task is a relevant tool for early diagnosis of MCI. The statistical analysis performed using the extracted data showed a significant difference in mean scores between MCI and control groups for the "animal" semantic fluency task. These results confirmed the relevance of this test in early diagnosis of MCI.

Keywords: mild cognitive impairment, verbal fluency, semantic fluency, assessment, diagnosis

20 pages et 3 pages d'annexe, 31 références bibliographiques.